eRD20: Developing Simulation and Analysis Tools for the EIC

Consortium

W. Armstrong (ANL), M. Asai (SLAC), E.-C. Aschenauer (BNL), D. Blyth (ANL), A. Bressan (Trieste), S. Chekanov (ANL), W. Deconinck (William & Mary), M. Diefenthaler (JLAB), A. Kiselev (BNL), D. Lawrence (JLAB), J. Lauret (BNL), C. Pinkenburg (BNL), S. Prestel (Lund), D. Romanov (JLAB), E. Saavedra (CERN), M. Ungaro (JLAN), D. Wright (SLAC)

Markus Diefenthaler, Alexander Kiselev

Jefferson Lab

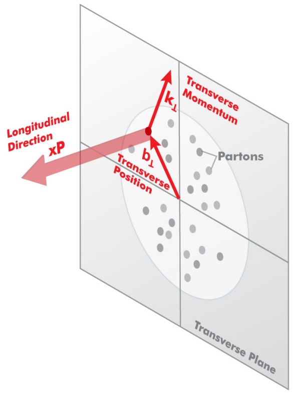
BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY | Office of Science

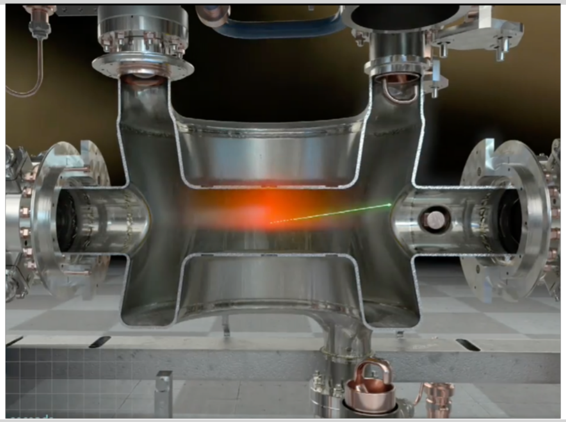
JSA

Scientific advances a new frontier in Nuclear Physics

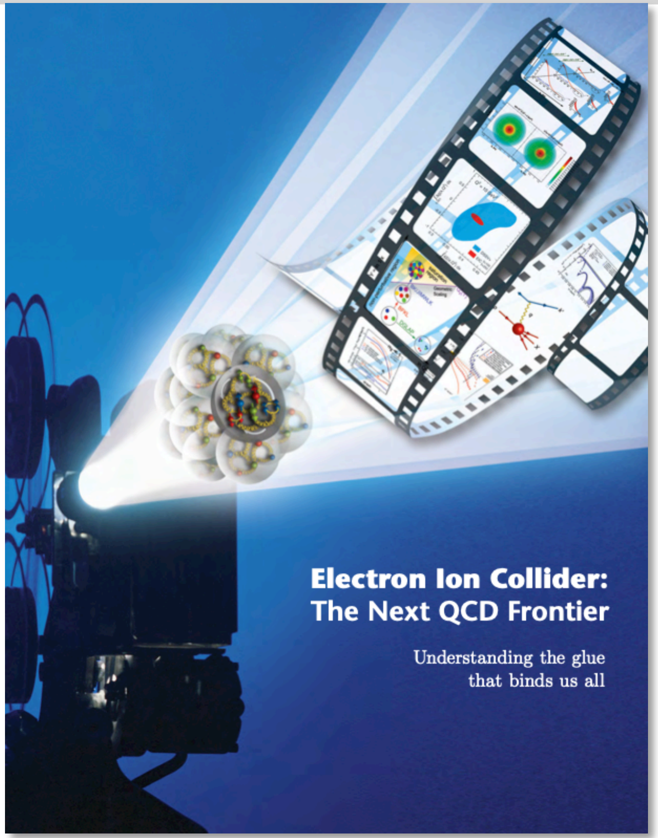
Quantum Chromodynamics



Accelerator technologies



Electron-Ion Collider: Steady advances in all of these areas will enable precision study of the nucleon and the nucleus at the scale of sea quarks and gluons.

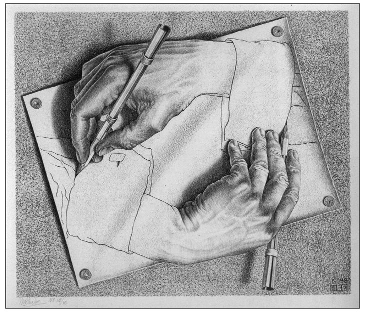


Detector technologies

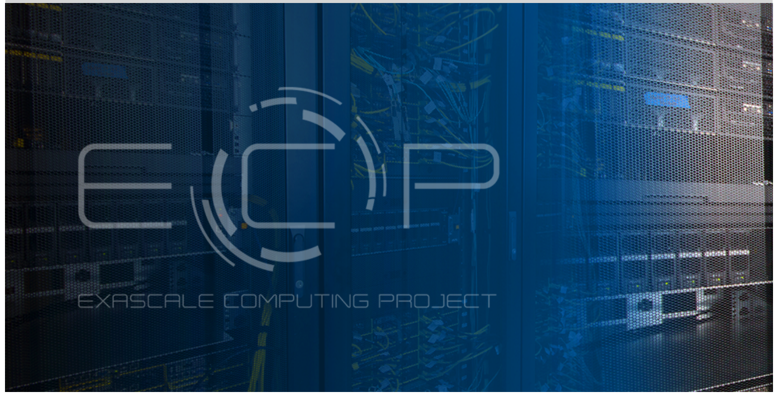
Electron-Ion Collider
Detector Requirements and
R&D Handbook

DRAFT 8 - December 15, 2018

Editors:
Alexander Kiselev
Thomas Ullrich



Computer technologies

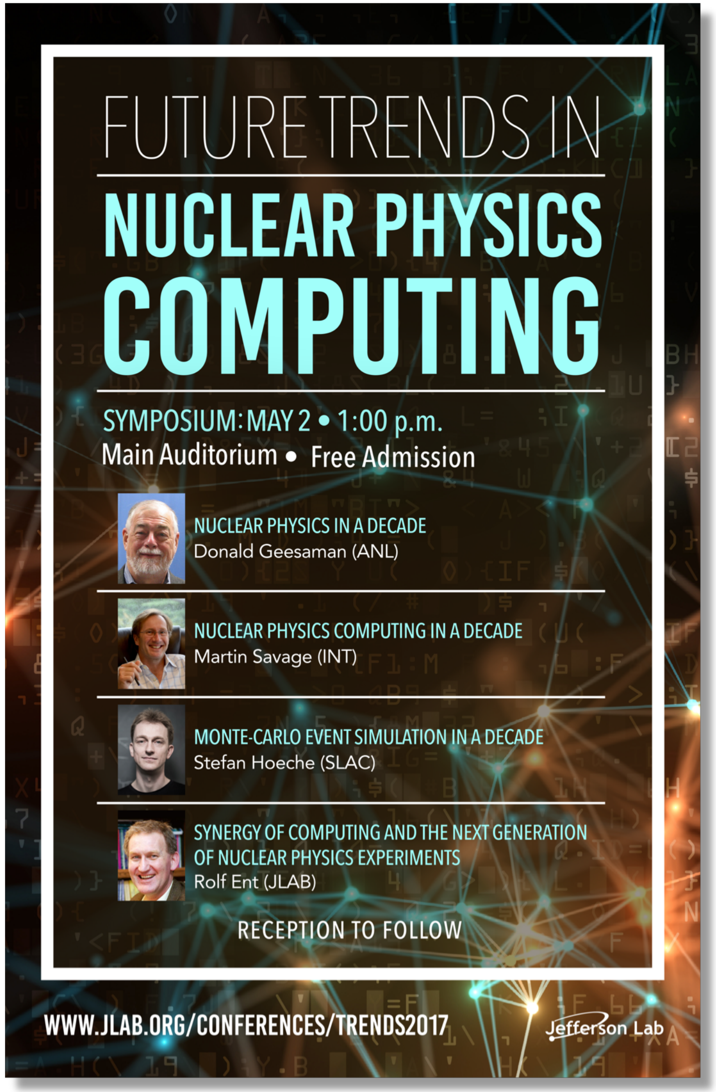


The central role of computing at the EIC

The purpose of computing is insight, not numbers.

Richard Hamming (1962)

Future Trends in Nuclear Physics Computing



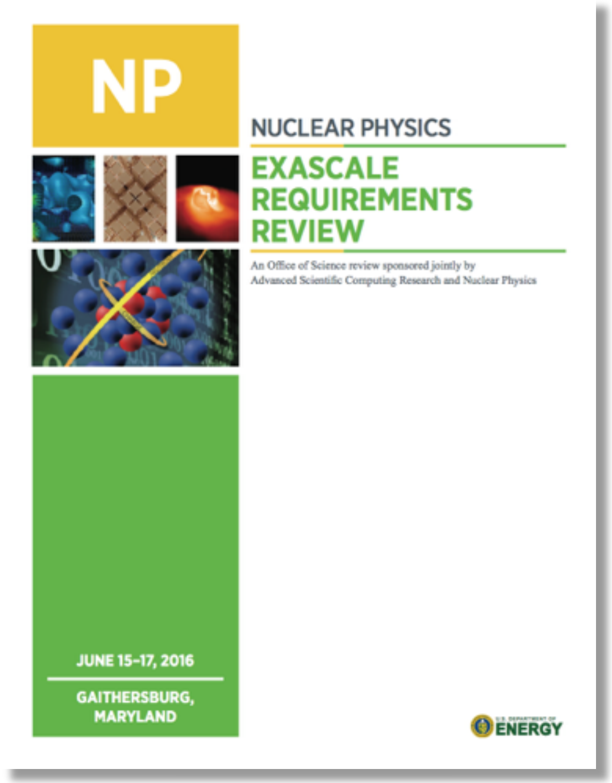
Donald Geesaman (ANL, former NSAC Chair) *"It will be **joint progress of theory and experiment** that moves us forward, not in one side alone"*

Martin Savage (INT) *"The next decade will be looked back upon as a **truly astonishing period in NP** and in our understanding of fundamental aspects of nature. This will be **made possible by advances in scientific computing** and in how the NP community organizes and collaborates, and how DOE and NSF supports this, to take full advantage of these advances."*

Implications of Exascale Computing

Past efforts in lattice QCD in collaboration with industry have driven development of new computing paradigms that benefit large scale computation. These capabilities underpin many important scientific challenges, e.g. studying climate and heat transport over the Earth.

The EIC will be the facility in the era of high precision QCD and the first NP facility in the **era of Exascale Computing**. This will affect the interplay of experiment, simulations, and theory profoundly and result in a new computing paradigm that can be applied to other fields of science and industry.



Petascale-capable systems at the beamline

- **unprecedented compute-detector integration**, extending work at LHCb
- requires fundamentally new and different algorithms
- computing model with machine learning at the trigger level and a compute-detector integration to deliver **analysis-ready data from the DAQ system**:
  - responsive calibrations in real time
  - real-time event reconstruction
  - physics analysis in real time

Strong connection with EIC Streaming Readout Consortium

A similar approach would allow **accelerator operations** to use real-time simulations and artificial intelligence / machine learning over operational parameters to tune the machine for performance.

Towards the next generation research model in Nuclear Physics

NP research model not changed for over 30 years
Science & Industry remarkable advances in computing & microelectronics

vision evolve & develop **NP research model** based on these advances

rethink

how measurements are compared to theory
- examine capabilities of event level analysis (**ELA**) taking the multi-dimensional challenges of NP fully into account

how experimental data are handled
- identify ways to speed up analysis in the context of **ELA**

how we read out detectors and assemble detector data
- investigate capabilities of streaming readout in view of **ELA**

IR design, EIC Detector R&D
Machine – Detector – Analysis Interface
EIC Streaming Readout and Software Consortia

goal

EIC Software Consortium

ESC members

ANL, BNL, CERN, JLAB, Lund, SLAC, Trieste, W&M

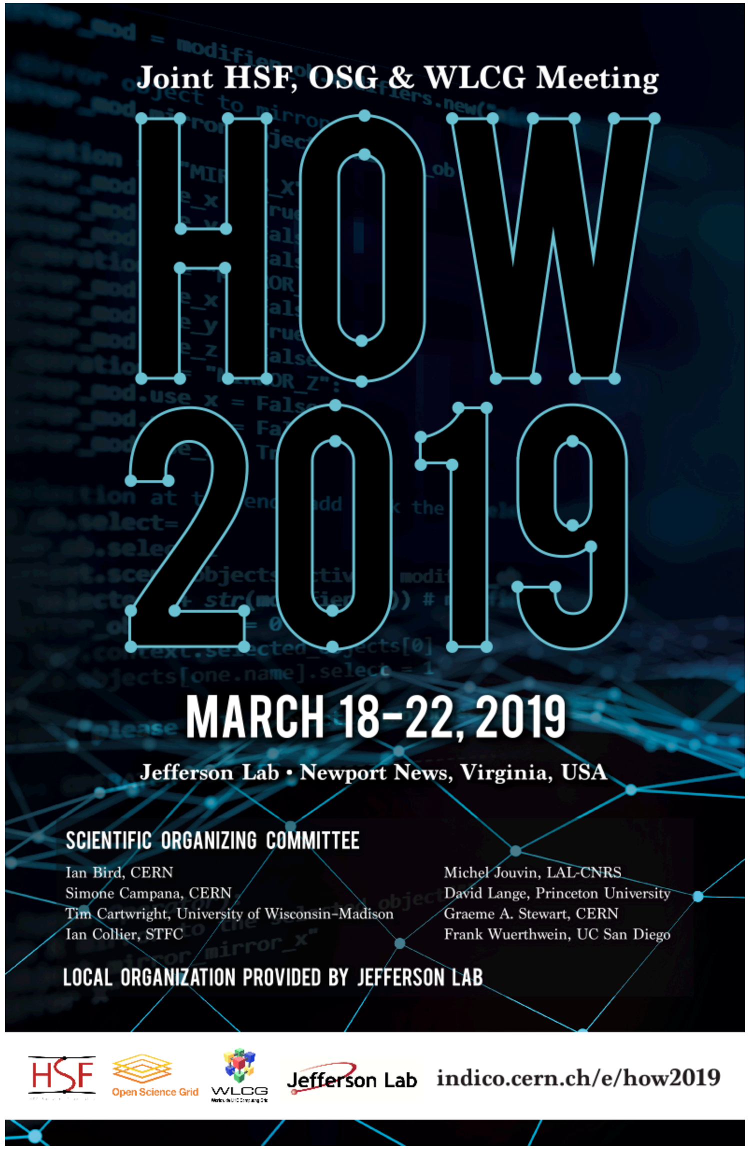
ESC goals and focus

- continue work on common interfaces (e.g., geometry, file formats, tracking)
- explore new avenues of software development (e.g., artificial intelligence)
- **reach out to the EIC community**
  - communicate present status of EIC software
  - bring existing EIC software to the end users
  - produce publicly available consensus-based documents on critical subjects
  - provide vision for the future

BROOKHAVEN NATIONAL LABORATORY Jefferson Lab

Collaboration

Geant4 International Collaboration	**Liaison** Makoto Asai (SLAC)
HEP Software Foundation	**Topics** AI/ML, real-time processing, reconstruction
MCnet	**Liaison** Stefan Prestel (Lund)
ROOT	**Liaison** Enric Saavedra (CERN)



Represent EIC community at HOW 2019

EIC User Group Software Working Group

Conveners

David Blyth

ANL Postdoctoral Appointee

EIC Software Consortium

Markus Diefenthaler

Jefferson Lab Staff Scientist

EIC Software Consortium

Charge The EICUG Software Working Group's initial focus will be on simulations of physics processes and detector response to enable quantitative assessment of measurement capabilities and their physics impact. This will be pursued in a manner that is accessible, consistent, and reproducible to the EICUG as a whole. It will embody simulations of all processes that make up the EIC science case as articulated in the White-paper. The Software working group is to engage with new major initiatives that aim to further develop the EIC science case, including for example the upcoming INT program(s), and is anticipated to play key roles also in the preparations for the EIC project(s) and its critical decisions. **The working group will build on the considerable progress made within the EIC Software Consortium (eRD20)** and other efforts. The evaluation or development of experiment-specific technologies, e.g. mass storage, clusters or other, are outside the initial scope of this working group until the actual experiment collaborations are formed. The working group will be open to all members of the EICUG to work on EICUG related software tasks. It will communicate via a new mailing list and organize regular online and in-person meetings that enable broad and active participation from within the EICUG as a whole.

EIC Software Consortium

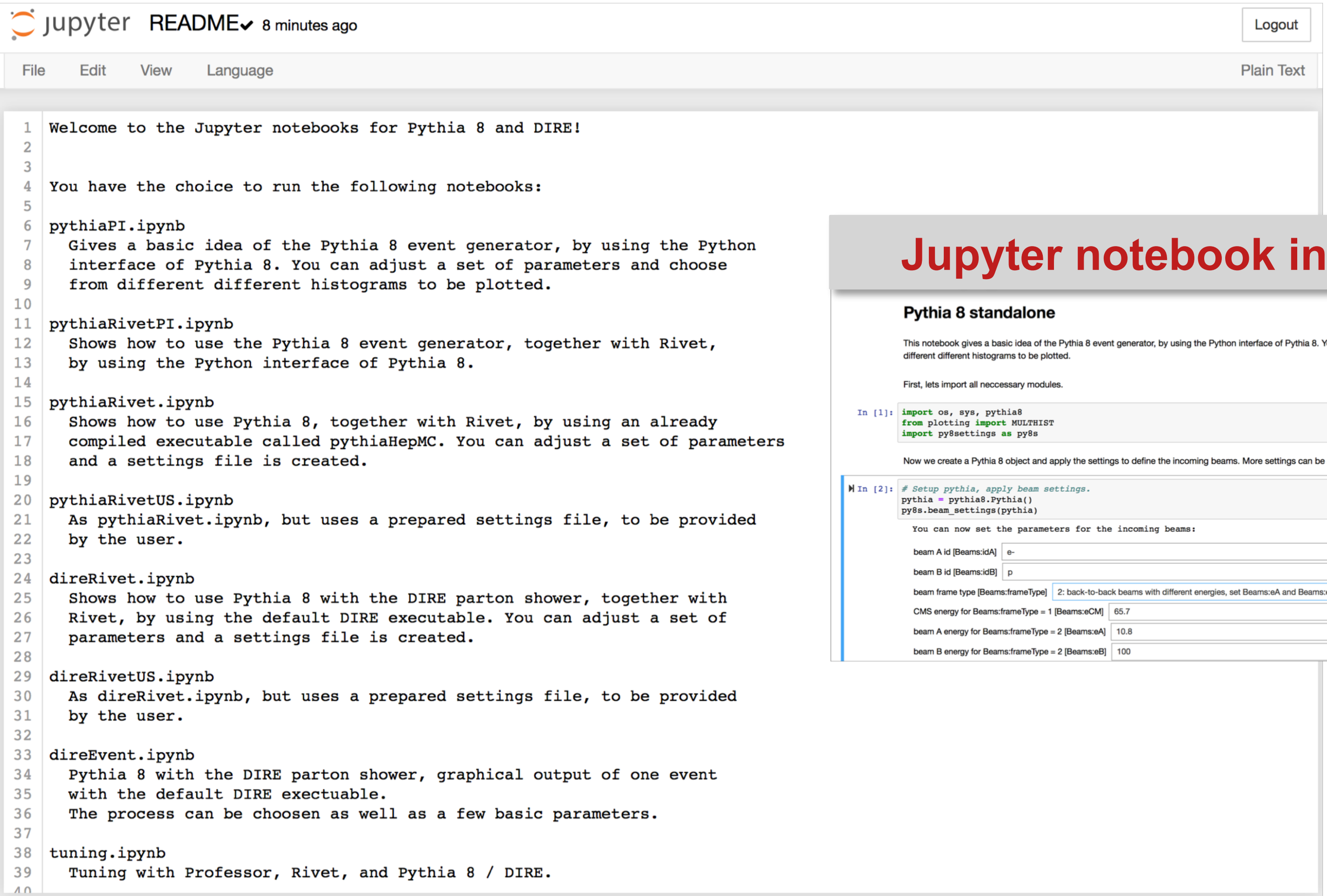
Status

Monte Carlo Initiative

- collaboration with MCnet
  - liaison: Stefan Prestel (Lund))
- reach out to EIC Community
  - work on online catalogue of MCEGs (in progress)
  - document on MCEG requirements (in progress)
- MCEG R&D:
  - containers and tutorials for EIC MCEGs (started)
  - library for QED radiative effects (in progress)

Example MCEG container

Container for Pythia8+DIRE by Nadine Fischer (Pythia)




```
Welcome to the Jupyter notebooks for Pythia 8 and DIRE!


You have the choice to run the following notebooks:

pythiaPI.ipynb
  Gives a basic idea of the Pythia 8 event generator, by using the Python
  interface of Pythia 8. You can adjust a set of parameters and choose
  from different different histograms to be plotted.

pythiaRivetPI.ipynb
  Shows how to use the Pythia 8 event generator, together with Rivet,
  by using the Python interface of Pythia 8.

pythiaRivet.ipynb
  Shows how to use Pythia 8, together with Rivet, by using an already
  compiled executable called pythiaHepMC. You can adjust a set of parameters
  and a settings file is created.

pythiaRivetUS.ipynb
  As pythiaRivet.ipynb, but uses a prepared settings file, to be provided
  by the user.

direRivet.ipynb
  Shows how to use Pythia 8 with the DIRE parton shower, together with
  Rivet, by using the default DIRE executable. You can adjust a set of
  parameters and a settings file is created.

direRivetUS.ipynb
  As direRivet.ipynb, but uses a prepared settings file, to be provided
  by the user.

direEvent.ipynb
  Pythia 8 with the DIRE parton shower, graphical output of one event
  with the default DIRE exectuable.
  The process can be choosen as well as a few basic parameters.

tuning.ipynb
  Tuning with Professor, Rivet, and Pythia 8 / DIRE.
```

Jupyter notebook interface

Pythia 8 standalone

This notebook gives a basic idea of the Pythia 8 event generator, by using the Python interface of Pythia 8. You can adjust a set of parameters and choose from different different histograms to be plotted.

First, lets import all neccessary modules.

```
In [1]: import os, sys, pythia8
        from plotting import MULTHIST
        import py8settings as py8s
```

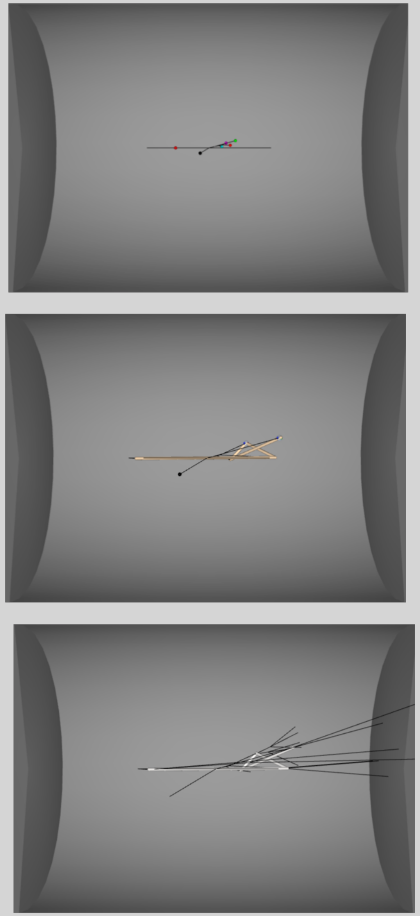
Now we create a Pythia 8 object and apply the settings to define the incoming beams. More settings can be adjusted later.

```
In [2]: # Setup pythia, apply beam settings.
        pythia = pythia8.Pythia()
        py8s.beam_settings(pythia)
```

You can now set the parameters for the incoming beams:

beam A id [Beams:idA]	e-
beam B id [Beams:idB]	p
beam frame type [Beams:frameType]	2: back-to-back beams with different energies, set Beams:eA and Beams:eB
CMS energy for Beams:frameType = 1 [Beams:eCM]	65.7
beam A energy for Beams:frameType = 2 [Beams:eA]	10.8
beam B energy for Beams:frameType = 2 [Beams:eB]	100

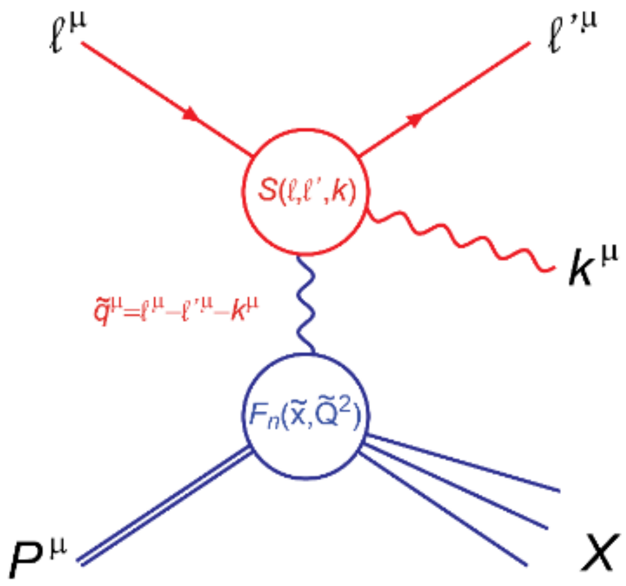
Visualization of ep collision



Radiative Effects and MCEG

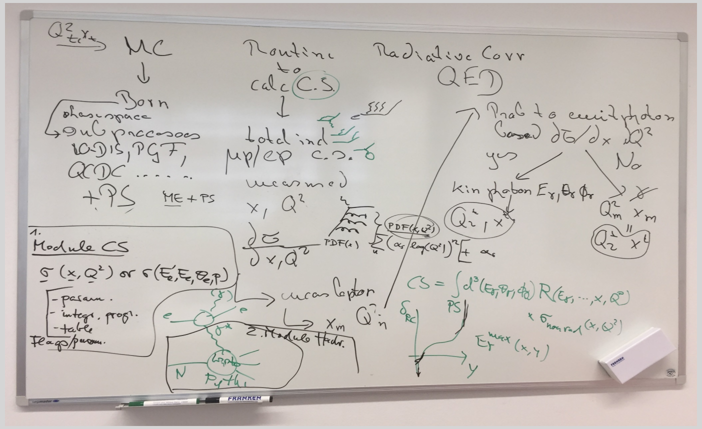
Radiative effects

- change kinematics on an event by event basis:
  - smearing of kinematic distributions
- change of virtual-photon direction:
  - false asymmetries in the azimuthal distribution of hadrons
- correction:
  - unfolding procedure, requires MCEG including radiative corrections / effects



ESC Radiative effects library

- Elke-Caroline Aschenauer, Andrea Bressan
- essential for high-precision measurements at the EIC
- collaboration with Hubert Spiesberger:
  - start back from HERACLES part of Djangoh
  - work on interface to PYTHIA6/8



MCEGs for future ep and eA facilites

Goals

- requirements for MCEGs for upcoming ep and eA measurements
- roadmap for developments of MCEGs for upcoming ep and eA measurements

Focus of this meeting will be on ep with addressing in particular

- status of ep and eA in general-purpose MCEG
- status of NLO simulations for ep/eA
- GPDs and TMDs in MC generators
- radiative / higher order QED effects in MCs and their treatment together with QCD effects

BROOKHAVEN NATIONAL LABORATORY Jefferson Lab

Detector Simulation

- **collaboration with Geant4**
  - liaison: Makoto Asai (SLAC)
- reach out to EIC Community
  - online catalogue of detector simulations (started)
- **Detector Simulation R&D**
  - containers and tutorials for EIC detector simulations (available)
  - coordinate input for Geant4 validation based on EIC physics list maintained by SLAC Geant4 group (work in progress)

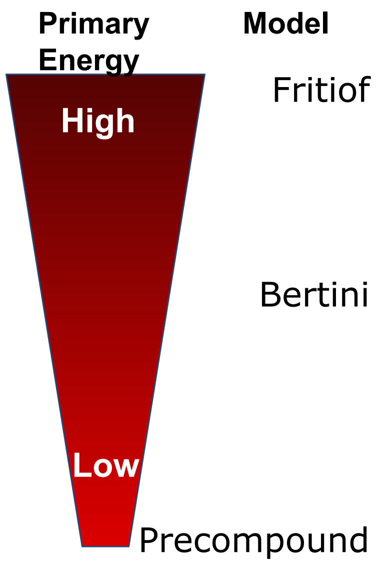
Technical forum on EIC as part of

09/23 – 09/27 Geant4 Collaboration Meeting (JLAB)

coordinate with EIC Detector R&D

EIC

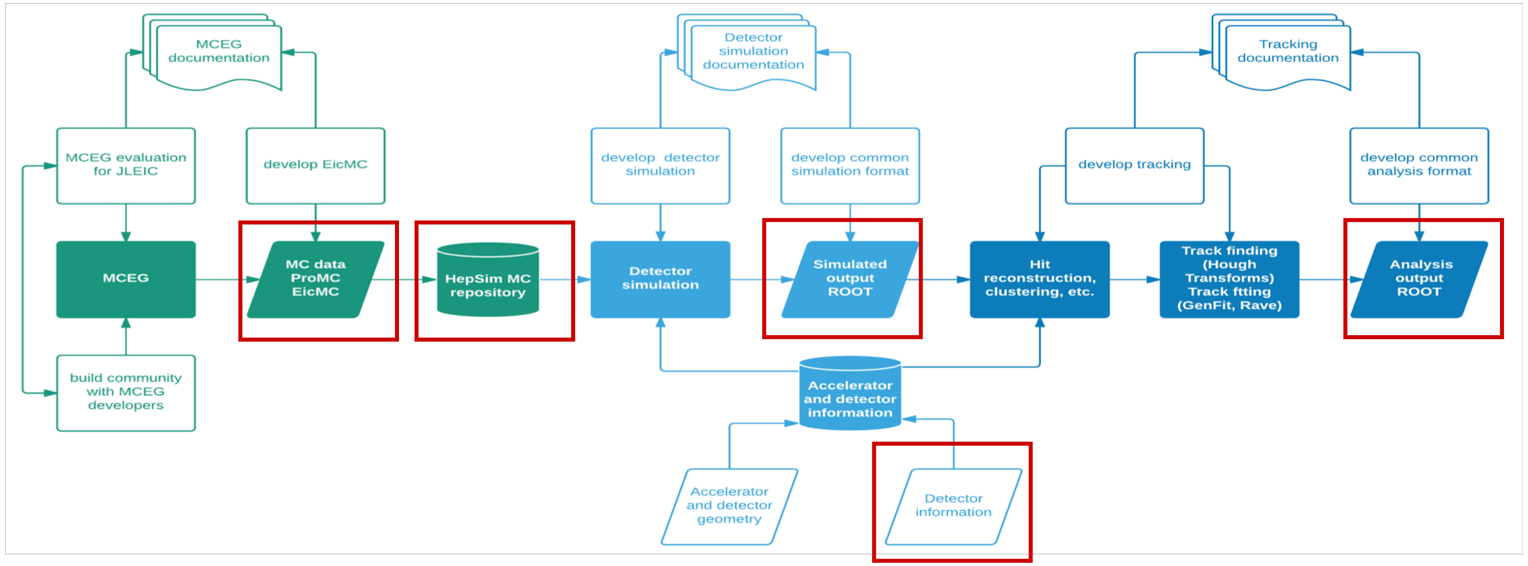
- energy range is different from LHC
- validation, tuning and extension including test beam studies



Common interfaces

Advice from ILC effort

- facilitate interoperability
- focus on exchange detector designs and data
  - get the event data model right and keep it open
  - pick a detector definition which is exchangeable



Norman Graf (SLAC)
“It’s very difficult to ~~herd cats~~ keep physicists from re-inventing the wheel and writing new software packages.“

BeAST container (EICroot)

ePHENIX container (fun4all)

MCEG

reconstruction

JLEIC container

TOPSiDE container

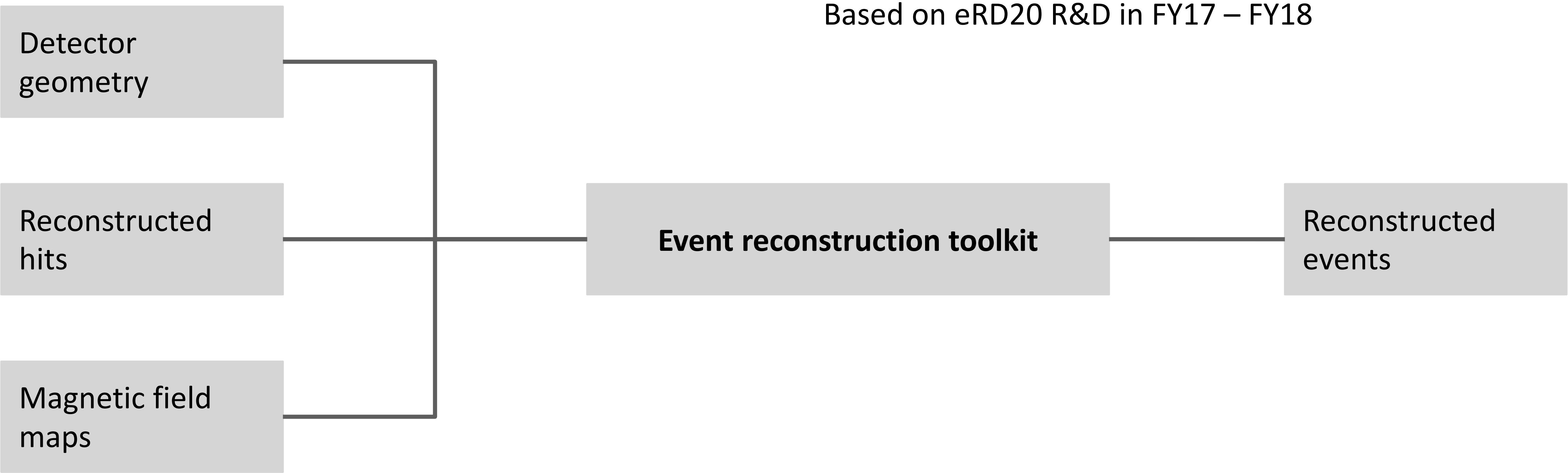
Community reference reconstruction

Charge *"The EICUG Software Working Group's initial focus will be on simulations of physics processes and detector response to enable* ***quantitative assessment of measurement capabilities and their physics impact****. This will be pursued in a manner that is* ***accessible****,* ***consistent****, and* ***reproducible*** *to the EICUG as a whole.*

Modular reconstruction based on EIC tracking tools (ANL, BNL, JLab)

- for interoperability of lab software tools
- for comparing / validating EIC results
- for testing new algorithms

Based on eRD20 R&D in FY17 – FY18



Components of event reconstruction

Reconstruction

- Calorimeter clustering
- Track reconstruction
- Vertex reconstruction

PID

- Lepton-hadron separation
- Lepton identification
- Hadron identification

Event kinematics

- Kinematics calculation

Components

- implementation of an algorithm
- developed by various scientists / groups

Work with EIC Detector R&D

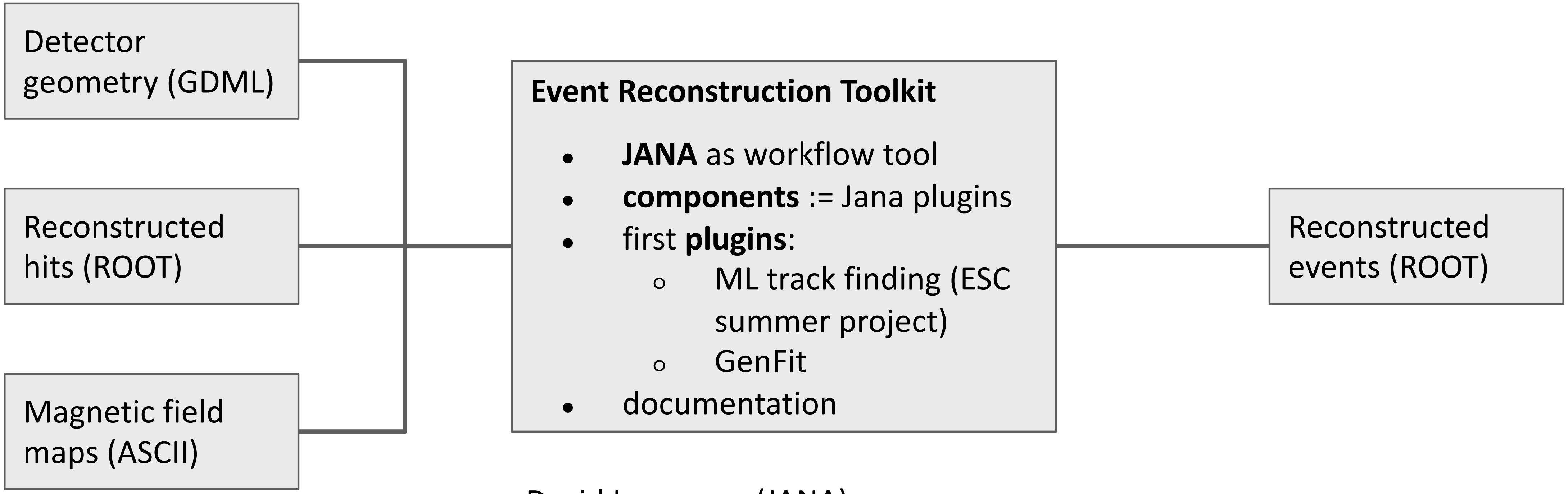
Benefits of community reference reconstruction

- **R&D on detector concepts**
  - baseline reconstruction
  - flexible reconstruction not optimized for / constrained by detector geometry

- **R&D on detectors**
  - test environment for R&D on new components

- **R&D for a community**
  - catalog of components
  - comparison of components

eJANA: Prototype for community reference reconstruction



David Lawrence (JANA)
Dmitry Romanov (eJANA)
Julia Furletova (Detector)

Event Data Model

We plan to encourage and support the free exchange of

- MCEG data
- simulated/reconstructed data
- reconstruction software

with collaborative development of an event data model at the EIC.

This model will be defined at several key points in the simulation/reconstruction process.

ProIO: Support of a collaborative data model

- In contrast to using C++ classes to form a data model, **ProIO** uses Google's Protocol Buffers to implement a **language-independent data model**
- This means that C++, Python, Java, Go, etc. are all first-class.
- Native Python code, e.g., can be `pip` installed in seconds on any OS to enable full read/write support of data (see py-proio)
- ProIO supports multiple compression algorithms, self description, and metadata

proio-org / **proio**

Unwatch 1 | Star 0 | Fork 0

Code | Issues 1 | Pull requests 0 | Projects 0 | Wiki | Insights | Settings

Branch: master | proio / model / eic / **eic.proto** | Find file | Copy path

decibelcooper Added language-specific options back into proto files — bf4bbdd 17 days ago

1 contributor

143 lines (124 sloc) | 4.02 KB — Raw | Blame | History

```
syntax = "proto2";
package proio.model.eic;
option go_package = "eic";
option java_package = "proio.model";
option java_outer_classname = "Eic";

////// TRUTH LEVEL DATA MODEL MESSAGES //////

message Particle {
    // ProIO entry identifiers that point to parent Particles
    repeated uint64 parent = 1;
    // ProIO entry identifiers that point to child Particles
    repeated uint64 child = 2;
    // PDG code
    optional sint32 pdg = 3;
    // position in mm and time in ns
    optional XYZTD vertex = 4;
    // momentum in GeV
    optional XYZF p = 9;
    // mass in GeV
    optional float mass = 10;
    // charge in units of e
    optional float charge = 7;
    optional XYZF spin = 8;

    // DEPRECATED
    optional XYZD p_double = 5;
    optional double mass_double = 6;
}

////// SIMULATION LEVEL DATA MODEL MESSAGES //////

message SimHit {
    // optional ID to circumvent volume lookup from gl
    // digitization
```


BROOKHAVEN NATIONAL LABORATORY

Jefferson Lab

AI/ML

Keynote

Advances in Trinity of AI: Data, Algorithms & Compute
Prof. Animashree Anandkumar (Caltech)

Foundations

DOE Scientific Machine Learning & AI Overview
Dr. Steven Lee (DOE ASCR)

Study of Neural Network Size Requirements for Approximating Functions Relevant to HEP

Jessica Stietzel (Notre Dame)

Deep learning for HEP/NP at NERSC

Dr. Wahid Bhimji (NERSC)

Applications

Tag Jet Identification Through the Use of Neural Networks
Anne-Katherine Burns (William & Mary)

Machine Intelligence Applications for Particle Physics at Fermilab
Dr. Aristeidis Tsaris (Fermilab)

Overview of Bayesian Optimization Applied to the GlueX case
Cristiano Fanelli (MIT)

2019 World Tour

- work on tutorials and example use cases in progress
- scheduling started for U.S. and Europe (Trieste)
- stay tuned for announcement of first dates

BROOKHAVEN NATIONAL LABORATORY Jefferson Lab

Markus Diefenthaler, Alexander Kiselev

Summary

mdiefent@jlab.org, ayk@bnl.gov

Sustained Progress

- work with EIC community
- key role in EIC Software Working Group

- work with scientific software community
- active role in software-related workshop organization
- arrange expert discussions, come to a consensus-based decisions, and document them in community documents

- take measures to prevent future EIC software divergence
- try to establish forward looking vision of EIC software